

Amélioration de la robustesse des bases POD

M. BERGMANN^{a,b}, E. LOMBARDI^{b,c}, A. IOLLO^{a,b}

a. Institut National de Recherche en Informatique et en Automatique (INRIA), équipe MC2, centre de Bordeaux Sud Ouest, 351, cours de la libération, 33405 Talence cedex, France

b. Institut de Mathématiques de Bordeaux (IMB), Université Bordeaux 1, UMR 5251 CNRS, 351, Cours de la Libération, 33405 Talence cedex, France

c. Dipartimento Ingegneria Aerospaziale, Università di Pisa, Via G. Caruso 8, 56122 Pisa, Italy

Résumé :

Cette étude concerne l'amélioration de la robustesse du sous-espace propre construit par Décomposition Orthogonale aux valeurs Propres (POD). Puisqu'une base POD est uniquement capable de donner une représentation optimale de l'énergie incluse dans la base de données, cette même base n'est pas adaptée pour représenter une dynamique d'écoulement engendrée avec d'autres paramètres d'entrée (paramètres de contrôle, nombre de Reynolds). L'objectif est donc de construire une base POD robuste capable de représenter tout un ensemble de dynamiques. Notre approche est basée sur une technique qui consiste à enrichir de façon itérative la base de donnée avec des réalisations calculées où l'erreur commise par le modèle réduit est la plus élevée ("Greedy method"). Des résultats issus du sillage d'un barreau dans un canal confiné seront présentés. Les paramètres d'entrée seront limités à un intervalle pour le nombre de Reynolds qui correspond au régime périodique 2D. On montrera alors qu'il est possible d'améliorer l'efficacité de la méthode "Greedy", en temps CPU en approximant l'erreur par une norme appropriée des résidus des équations de Navier-Stokes, ainsi qu'en terme d'approximation sur tout l'intervalle considéré.

Abstract :

This study focuses on improving the robustness of the functional subspace built using Proper Orthogonal Decomposition (POD). Since a POD basis is able to give an optimal representation of the kinetic energy included in the snapshots database generated with some given input parameters, this same basis is not adapted to represent flow dynamics generated with other input parameters. Our aim is thus to build a robust POD basis that accurately represents the solution over a desired input parameter subspace by enlarging the statistical database. We present a systematic method to sample the input parameter subspace. The basic idea is to add to the existing database, snapshots of the solution for which the POD approximation error is maximal. Then the new POD basis is obtained from the set of snapshots corresponding to such different dynamics. This is the so called Greedy sampling. The approach we follow is similar : it is based on finding the centroid of a region around the point where an estimate of the POD approximation error is maximal. We show a priori numerical evidence that the Navier-Stokes residuals are a reliable estimate of the POD approximation error. Results relative to a 2D confined square cylinder wake flow are presented. The input parameter subspace is represented by an interval of Reynolds numbers that corresponds to periodical laminar flows. We show that a judicious choice of the sampling Reynolds numbers leads to a POD basis that minimizes the average approximation error on the chosen interval.

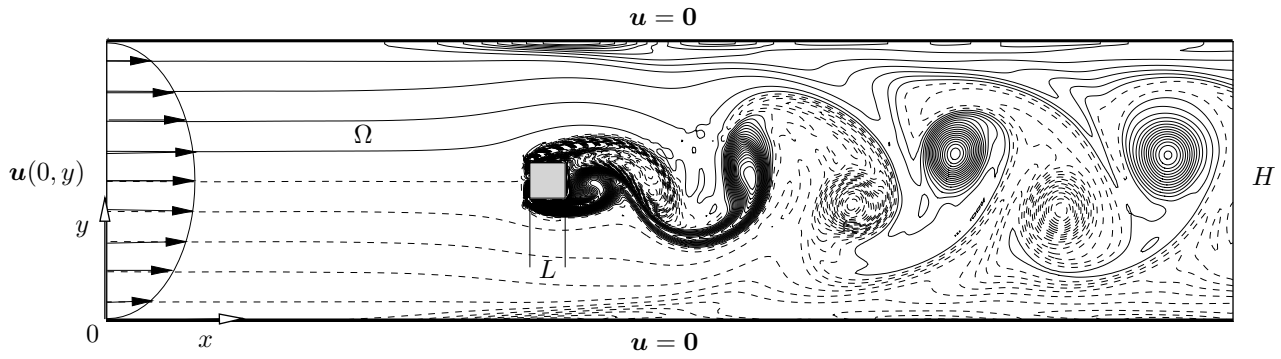
Mots clefs : modèles réduits, POD, échantillonnage, robustesse

1 Introduction

1.1 Réduction de modèle et expertise

Un modèle mathématique qui gouverne l'évolution d'un système complexe doit être maniable ou utilisable. La réduction et l'expertise de ces systèmes sont alors nécessaires. La réduction consiste à trouver un système plus "simple" ayant, au moins dans une gamme de paramètres donnée, le même comportement que le système complexe détaillé. L'expertise consiste, quant à elle, à déterminer les bonnes gammes et échelles des paramètres du système. Ces deux activités sont fortement liées entre elles, et négliger l'expertise peut amener à des modèles dont la solution est erronée. L'expertise peut par exemple consister à trouver une région de confiance pour le modèle réduit [?, ?]. La solution idéale au problème de la réduction est de trouver une solution analytique au modèle complet considéré. Malheureusement, ceci s'avère tout à fait impossible en général. Une alternative est alors de chercher des modèles approchés dont on puisse maîtriser la solution et dont le comportement soit proche de celui du système détaillé.

Il y a, entre autres, deux approches classiques adaptées à la réduction : les méthodes asymptotiques et d'in-

FIG. 1 – Configuration d'écoulement et champ de vorticité à $Re = 200$.

variance et les méthodes de projection. Les méthodes asymptotiques et d'invariance consistent à simplifier le modèle et ses équations. Les méthodes de projection consistent, quant à elles, à simplifier l'espace de phase dans lequel varient les paramètres, en autorisant les grandeurs dépendantes à évoluer que dans un espace de dimension finie. Dans un certain sens, une discrétisation éléments finis peut déjà être considérée comme une réduction de modèle. Parmi les nombreuses méthodes de projection qui existent, nous pouvons citer les bases de Lagrange, les bases d'Hermite, les bases de Taylor, la Décomposition Orthogonale aux valeurs Propres (POD, *Proper Orthogonal Decomposition*) [?, ?], les "balanced truncation" [?, ?], la décomposition sur des modes globaux de stabilité [?], ou tout autre décomposition Galerkin [?] *etc.* En raison de l'optimalité énergétique de son sous-espace propre, la POD est la méthode de réduction la plus utilisée en mécanique des fluides. Même si la POD a été utilisée initialement pour extraire les structures cohérentes des écoulements turbulents [?], elle a été utilisée plus récemment comme une technique simple et efficace de réduction pour différents problèmes de contrôle d'écoulement [?, ?, ?, ?, ?, ?]. Malheureusement la plupart de ces études ne se sont pas intéressées au point essentiel qu'est l'expertise des modèles réduits construits par POD (POD ROM). En effet, il a été démontré [?] que la solution d'un POD ROM pouvait être grandement détériorée en utilisant des paramètres du système assez différents de ceux utilisés pour générer le modèle.

Pour des problèmes de contrôle d'écoulement, le modèle doit pouvoir représenter le chemin dans l'espace des paramètres de contrôle qui mène de la condition initiale jusqu'à l'optimum recherché. Malheureusement, l'expertise montre que l'efficacité des POD ROM à correctement approximer le modèle détaillé est limitée à un voisinage proche des paramètres utilisés pour dériver le POD ROM.

La première classe de méthode d'expertise est d'actualiser de façon itérative le modèle pendant le processus d'optimisation. Une méthode efficace est la méthode à région de confiance couplée à la POD (TRPOD, *trust region POD*). La seconde classe de méthode d'expertise consiste à déterminer une base dite généralisée qui représente fidèlement une plus large gamme de dynamiques, sans actualisation de la base. Par exemple, la base POD peut être calculée à partir de réalisations obtenues en utilisant une loi de contrôle "chirp" qui balaye une large gamme de dynamiques [?, ?, ?]. Une fois la base déterminée, l'optimisation est alors très rapide car elle ne nécessite alors aucune résolution du modèle détaillé Navier-Stokes. Malheureusement, cette méthode dépend largement de la loi de contrôle "chirp" utilisée pour calculer la base POD. Le contrôle est alors sous optimal. Cette méthode nécessite alors une grande part d'empirisme. Dans le même esprit, la base peut être calculée en mélangeant des réalisations calculées pour différents paramètres de contrôle. Par exemple, l'échantillonnage aléatoire, l'échantillonnage orthogonal, "*Latin Hypercube Sampling*" (LHS), ou la "*Centroidal Voronoi Tessellation*" (CVT, [?, ?]) peuvent être utilisés. En dépit de l'efficacité de la CVT, il est nécessaire de fixer le nombre de points d'échantillonnage. Ces méthodes ne sont donc pas optimales dans le sens où il n'est pas possible de contrôler la précision des POD ROM à approximer le modèle détaillé. L'objectif de cette étude est alors de déterminer de façon itérative un échantillonnage optimal du sous-espace des paramètres d'entrée qui peut être arrêté quand un seuil de précision est atteint.

1.2 Configuration d'étude

Notre configuration est l'écoulement laminaire de sillage d'un cylindre de section carrée placé dans un canal 2D (voir figure ??). La condition d'entrée présente un profil de poiseuille. Des conditions de non-glissement sont imposées sur les murs parallèles (haut et bas). Le ratio de blocage est $L/H = 1/8$. Les détails des méthodes numériques utilisées sont reportés dans la Réf [?]. Cet écoulement est intéressant car il présente décollement de la couche limite, sillage et interaction avec les murs du canal.

Après avoir brièvement introduit la POD (§??), nous présenterons une méthode simple pour échantillonner le sous espace des paramètres d'entrée du système afin d'obtenir une base POD robuste (§??).

2 Décomposition Orthogonal aux valeurs Propres (POD)

La POD consiste à décomposer les champs physiques de la manière suivante [?] :

$$\bar{U}(\mathbf{x}, t) = \sum_{n=1}^{N_r} a_n(t) \Phi_n(\mathbf{x}), \quad (1)$$

où N_r est le nombre de modes retenus. La réduction consiste à trouver le plus petit N_r possible tout en conservant en degrés d'approximation raisonnable. La base POD $\Phi_n(\mathbf{x})$ est calculer à l'aide de la méthode des snapshots [?]. Les coefficients $a_n(t)$ sont solution d'un modèle réduit obtenu à l'aide d'une projection de Galerkin du modèle détaillé sur la base POD. Cependant, nous nous limiterons ici à l'étude de la robustesse de la base, et non du modèle (travail à venir), et donc les coefficients temporels $a_n(t)$ seront évalués comme étant la projection de la solution numérique des équations de Navier-Stokes $U(\mathbf{x}, t)$ sur la base POD $\Phi_n(\mathbf{x})$, i.e.

$$a_n(t) = \int_{\Omega} U(\mathbf{x}, t) \Phi_n(\mathbf{x}) d\mathbf{x}. \quad (2)$$

Afin d'évaluer les résidus de équations de Navier-Stokes (voir par la suite), la base POD est étendue au champ de pression. Les réalisations de l'écoulement seront alors $U(\mathbf{x}, t_i) = (u(\mathbf{x}, t_i), v(\mathbf{x}, t_i), p(\mathbf{x}, t_i))^T$.

3 Echantillonnage des paramètres d'entrée

Les fonctions de base POD sont optimales pour donner une représentation de l'énergie cinétique incluse dans la base de données. Habituellement, cette base de données est collectée à partir d'un écoulement généré avec des paramètres d'entrée du système fixés. La même base n'est alors pas optimale pour représenter les caractéristiques d'écoulements générés avec d'autres paramètres d'entrée (voir [?, ?, ?]). L'objectif de cette section est de donner un critère efficace pour échantillonner le sous espace des paramètres d'entrée pour améliorer la robustesse des fonctions de base POD. Pour des raisons de simplicité, le sous-espace des paramètres d'entrée est limité à un intervalle pour le nombre de Reynolds $\mathcal{I} = [Re_L, Re_R]$ avec $Re_L = 70$ et $Re_R = 180$. Ces valeurs correspondent approximativement aux limites basse et haute du régime périodique 2D pour l'écoulement considéré (§??). Numériquement, \mathcal{I} est discrétisé avec $\Delta Re = 5$, et est noté \mathcal{I}_h . Tous les concepts introduits dans cette étude seront facilement généralisables aux sous-espaces de contrôle de dimensions élevées comme dans la Ref. [?].

Dans le but d'améliorer le sous-espace propre POD nous voulons augmenter la taille de la base de données en ajoutant de manière itérative des ensembles de réalisations qui correspondent à différents nombres de Reynolds $Re_i \in \mathcal{I}$. Soit $U^{[Re_1, \dots, Re_N]}$ la base de données collectée avec N ensembles de réalisations pour Re_1, \dots, Re_N , où le paramètre N est fixé en fonction de la précision souhaitée. La question est alors comment choisir ces nombres de Reynolds pour calculer une base POD robuste de telle sorte que la reconstruction $\bar{U}(\mathbf{x}, t)$ (voir Eq. (??)) soit aussi proche que possible de la solution numérique du modèle détaillé de Navier-Stokes, $U(\mathbf{x}, t)$, pour tous les nombres de Reynolds inclus dans \mathcal{I}_h . Nous utiliserons toujours $N_r = 50$ fonctions de base, même lorsque $N > 1$ nombres de Reynolds seront considérés.

Deux classes de méthodes d'échantillonnage POD sont couramment utilisées. La première est basée sur la "Centroidal Voronoi Tessellations" (CVT, voir [?, ?, ?]) qui peut être calculée avec un algorithme de Lloyd [?]. L'inconvénient est qu'il faut fixer le nombre de points d'échantillonnage, et il n'est donc pas possible de contrôler la précision souhaitée. La seconde est basée sur des "Greedy sampling" (voir [?]). Ceci est un échantillonnage itératif dans lequel il est possible d'arrêter l'algorithme lorsqu'un seuil de précision désiré est atteint. Dans cette étude, nous présentons une méthode hybride qui utilise les avantages des deux méthodes décrites ci-dessus.

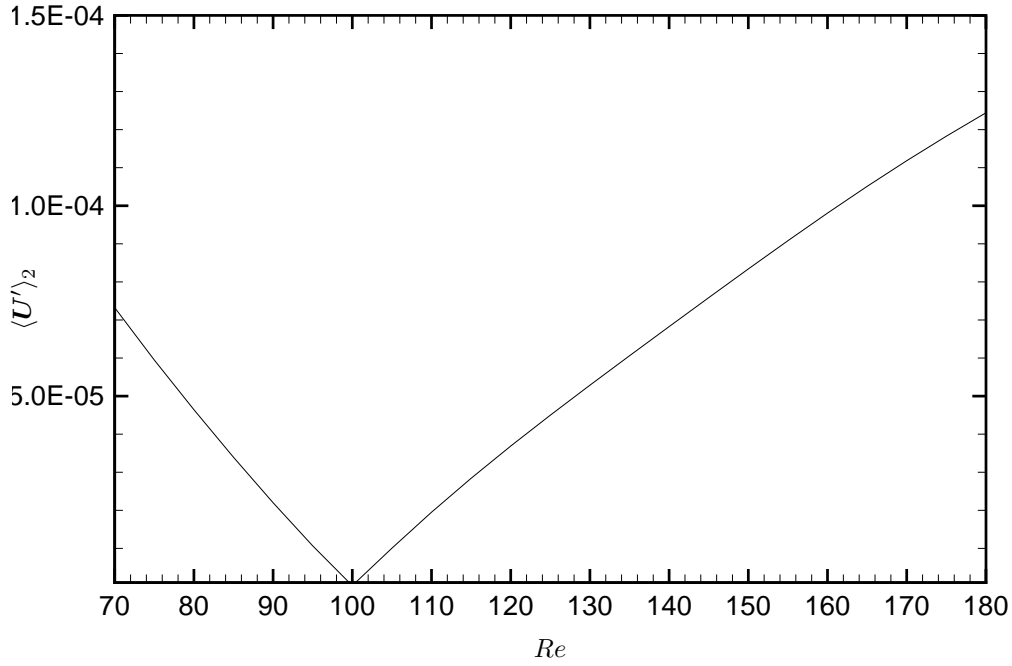
Pour illustrer notre méthode d'échantillonnage nous considérons une base de donnée initiale $U^{[Re_1]}$ composée de N_s réalisations collectées à $Re_1 = 100$. Nous utiliserons également toujours $N_s = 200$, et donc la base de données $U^{[Re_1, \dots, Re_k]}$ est composée de $k \times N_s$ réalisations. Comme nous pouvons voir sur la figure ??, l'erreur définie par

$$\langle U' \rangle_2 = \int_T \|U'(\mathbf{x}, t)\|_2 dt, \quad (3)$$

où $U'(\mathbf{x}, t) = U(\mathbf{x}, t) - \bar{U}(\mathbf{x}, t)$ représente les échelle non calculées, croit avec la distance $|Re - Re_1|$. La méthode "Greedy" consiste à échantillonner le sous espace des paramètres d'entrée où l'erreur est maximale, soit

$$Re_k^{GRE} = \arg \max_{Re} \langle U' \rangle_2. \quad (4)$$

Par exemple $Re_2^{GRE} = 180$ sur la figure ?? . Pour réduire les coûts de calcul CPU, il est possible d'utiliser les résidus de l'opérateur de Navier-Stokes évalué avec la reconstruction POD, \bar{U} (voir Eq. (??)), notée $\mathcal{R}_M(\bar{U})$,

FIG. 2 – Evolution de l'erreur $\langle U' \rangle_2$ en fonction du nombre de Reynolds.

au lieu d'utiliser l'erreur (??) (voir [?]). En utilisant cette approximation, les nouveaux points $Re_k, k = 2, \dots, N$ peuvent être calculés comme étant les centres de la région autour du point "Greedy" (CGR) \mathcal{I}_k^{GRE}

$$Re_k^{CGR} = \frac{\int_{\mathcal{I}_k^{GRE}} Re \rho(Re) dRe}{\int_{\mathcal{I}_k^{GRE}} \rho(Re) dRe}, \quad \text{with} \quad \rho(Re) = \langle \mathcal{R}_M(\bar{U}) \rangle_2, \quad (5)$$

où la région autour des points de "Greedy" est

1. centrée autour de la valeur "Greedy" Re_K^{GRE} ,
2. avec un rayon égale à $\min_{i \in [1; N-1]} |Re_K^{GRE} - Re_i|$,
3. restreint à $\mathcal{I}_K^{GRE} \subset \mathcal{I}$.

Ce critère peut facilement être généralisé à des sous-espaces de paramètres d'entrée de dimension supérieure à 1. Par exemple, les régions autour des points "Greedy" seraient des disques, sphères et hyper sphères pour des dimensions 2, 3 et 4 respectivement.

Nous pouvons voir sur le tableau ?? que ce critère permet une diminution importante de l'erreur moyenne, notée \mathcal{J} , évaluée sur tout l'intervalle \mathcal{I} , i.e.

$$\mathcal{J} = \int_{\mathcal{I}} \langle U' \rangle_2^2 dRe. \quad (6)$$

k	1	2	3	4	5
Re_k^{CGR}	100	152	81	170	126
\mathcal{J}_k^{CGR}	$60 \cdot 10^{-4}$	$16 \cdot 10^{-4}$	$7.7 \cdot 10^{-4}$	$3.6 \cdot 10^{-4}$	$3.0 \cdot 10^{-4}$
Re_k^{GRE}	100	180	70	140	80
\mathcal{J}_k^{GRE}	$60 \cdot 19.2^{-4}$	$19.2 \cdot 10^{-4}$	$12.9 \cdot 10^{-4}$	$6.1 \cdot 10^{-4}$	$4.5 \cdot 10^{-4}$

TAB. 1 – Evolution de l'erreur moyenne (??) en fonction du nombre de points d'échantillonnage.

Ce critère CGR, défini *a priori*, donne des points plus efficace que ceux de la méthode "Greedy" (voir figure ??). De plus, les points CGR sont très proches des points optimaux, définis *a posteriori*.

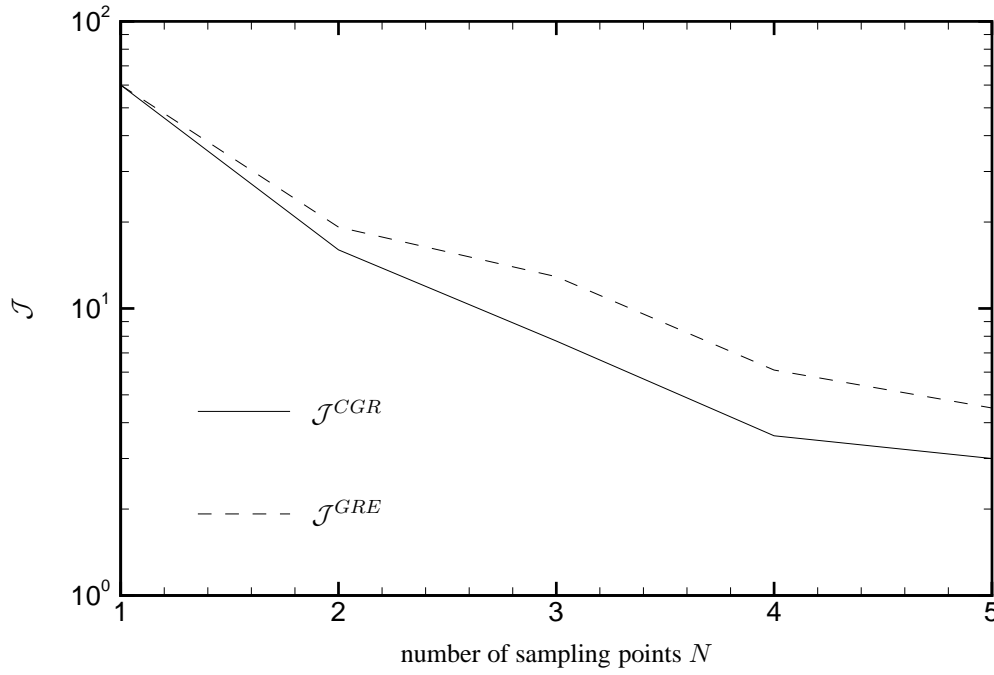


FIG. 3 – Evolution de l'erreur moyenne \mathcal{J} en fonction du nombre de points d'échantillonnage.

4 Conclusions

Actuellement, malgré les progrès faits en informatique et en algorithmique, il est encore difficilement concevable de faire du contrôle d'écoulement sans avoir recours à des modèles réduits. L'inconvénient majeur de la POD est que les bases sont uniquement capables de restituer l'information contenu dans la base de donnée. Il est alors nécessaire d'échantillonner intelligemment l'espace des paramètres de contrôle que l'on souhaite étudier. Dans cette étude, nous avons présenté une technique efficace d'échantillonnage. Premièrement, nous avons utilisé la norme L_2 de l'opérateur de Navier-Stokes évalués avec les champs POD pour approcher l'erreur d'approximation commise avec la base POD. Nous avons ensuite proposé un échantillonnage qui couple les idées de "Greedy method" et de "Centroidal Voronoi Tessalation". Cet échantillonnage est donc rapide (puisqu'il est uniquement nécessaire d'évaluer des résidus) et précis (il peut être montré que cet échantillonnage est proche de l'échantillonnage optimal). Dans la suite de cette étude, nous avons prévu de prendre en compte le modèle réduit. Ce travail est en cours.